TIME'S ARROW FROM THE MULTIVERSE POINT OF VIEW

MARTIN TAMM

Abstract. In this paper I suggest a possible explanation for the asymmetry of time. In the models that I study, the dynamical laws and the boundary conditions are completely symmetric, but the behaviour of time is not. The underlying mechanism is a spontaneously broken symmetry on the micro-level which is closely related to the idea of multiple histories in quantum mechanics. The situations that I will discuss are very simple and could even in a sense be called classical, but the character of the mechanism is so general that the results are likely to carry over to more complicated cases. Since the computational difficulties are enormous, I mainly use heuristic methods and computer computations to exploit these ideas.

Author keywords: multiverse, time's arrow, cosmology, entropy

Martin Tamm,
Dept of Mathematics,
Univ. of Stockholm,
S-106 91 Stockholm,
Sweden.
matamm@math.su.se

1. Introduktion

To explain the second law of thermodynamics and the non-symmetric behaviour of time is a fundamental problem in contemporary physics. In any situation known to us, the direction towards the future differs from the direction towards the past by the property that in the former entropy increases, but in the latter it decreases. The question is: What causes this asymmetric behaviour?

It may very well be that the problem of time is too complex to have a simple solution in the usual physical or mathematical sense of the word. But it can still be that we can clarify the problem in a convincing way by constructing models of the universe where the reasons for the break of time-symmetry can be seen clearly. A remote analogy might be the following: For more than 2000 years, mathematics was plagued by the problem with Euclid's fifth Postulate. Enormous efforts were made to prove it from the other Postulates, but once the right strategy was found (in this case to construct different models for geometries satisfying the other Postulates), the dispute could quickly be brought to an end.

In this paper, a central theme is the concept of multiple histories. The idea to use this concept to explain the asymmetry of time is not new, but my ambition in this paper is not so much to present new physics as to try to find a way out of

the confusion. Hence, I have tried to make the frame-work as simple as possible in order to make the fundamental problem come out more clearly.

A part of the problem is that we inevitably have to deal with two different view-points; The microscopic and the macroscopic perspectives. In this paper, this double perspective is always present. After a short analysis of the problem and a presentation of the multiverse from my point of view in Sections 2 and 3, I turn in Section 4 to the dynamical laws. In Section 5, I then give a very simple example of a (microscopic) dynamical principle which from a deterministic point of view is completely time-symmetric, but which from the multiverse point of view generates a break of symmetry. In Sections 6, 7 and 8 I then turn to the opposite (cosmological) perspective. Using heuristic methods I argue that a dynamical principle which generates multiple histories on the micro-level is also likely to generate a global broken symmetry of time. In Section 9 I then, using ideas that are very close to the ones in Section 5, try to model this behaviour in an extremely small multiverse containing only three particles. This model is at least within reach for computational methods. In Section 10, I conclude by making some remarks.

The reader may note one peculiar fact in the following: Although in a way the omnipresent concept is entropy, not a single formula or computation containing entropy is included. This is in the same time a problem and the motivation for this work. It is a problem since computing the entropy is really what I would like to do. In a better world than ours, the computational difficulties of this paper would perhaps be treated by brute force using computers. As it is, the computations in the following are, in their present form, both in a sense too difficult to be handled by computers and too simple to be satisfactory. But still, we have to make use of the possibilities we have at hand to reach understanding.

2. The riddle of time

Although the cosmological perspective on the issue of entropy in a sense goes back to Bolzmann, the starting-point for the modern development is Gold's article [3] in 1962. In this paper, Gold argues that the second law of thermodynamics might be a consequence of an expanding universe. A natural consequence of this idea is that in a contracting universe, entropy will decrease and we are therefore naturally lead to consider a completely symmetric model for the universe with low entropy in both ends.

Today however, few cosmologists support this view of a symmetric behaviour of the entropy, and hence most efforts since Gold have concentrated on trying to understand why the behaviour is non-symmetric. The main strategies are:

- We can try to explain the arrow of time by making appeal to some kind of lack of symmetry in the boundary conditions of the universe.
- We can try to explain the arrow of time by making appeal to some kind of lack of symmetry in the laws of physics themselves.

Clearly, if we somehow suppose that the universe immediately after the Big Bang was very smooth and ordered, and that the ultimate future of the universe (whether viewed as a Big Crunch or as an eternal expansion) will be very disordered, then the increase of entropy in-between may appear quite natural. But this still leaves us with the question why the boundary conditions should be so highly asymmetric.

Steven Hawking, in his early attempts to understand the arrow of time in terms of his "no boundary condition approach", adopted Gold's view. Later, after criticism from Don Page and Raymond Laflamme, he abandoned this idea and extended his analysis of the no boundary condition to conclude that there could be two types of behaviour of the universe close to the end-points, one which fits nicely with a low-entropy Big Bang, and another one which fits with a high-entropy Big Crunch (see [4]). It can be argued that the real problem is to give a satisfactory explanation of how these different boundary behaviours are coupled with each other. Nevertheless, it may be that Hawking's analysis is an important part of the answer.

An example of the another kind of explanation is given by Penrose's suggestion that the growth of entropy is connected with the Weyl tensor [5]. There are also attempts to connect the arrow of time with the apparently time-asymmetric behaviour of the K-meson. If we adopt this view, then the dynamical laws become explicitly non-symmetric. The problem in this case is that it seems very difficult to connect this with the second law of thermodynamics which is basically a macroscopic, classical phenomenon.

The point that I want to make is that it may in fact be that neither of the two strategies above is actually necessary to explain the asymmetry of time. I will argue that if we adopt the multiverse point of view, the ideas of symmetric laws and of a symmetry between the initial state and the end-state of our universe are very well compatible with the kind of asymmetric behaviour of time that we can observe. The underlying principle for this phenomenon can be described as follows: When we replace a symmetric deterministic dynamical law by a non-deterministic one, a break of symmetry occurs which generates multiple histories in one of the two possible directions of time. Although the situations which I will discuss in this paper are very simple and essentially classical, the mechanism in it self appears to be of such a general character that it may very well carry over to more complex geometric and quantum mechanical situations.

3. The multiverse view-point

According to the common Copenhagen interpretation of quantum mechanics, time development is deterministic but the wave-function does not determine the exact position and momentum of a given particle. Any measurement effects the particle (or the associated wave-function) in such a way that the particle, after the measurement, will be in an eigenstate of the corresponding dynamical variable. This view-point has been remarkably effective for computing all kinds of results.

However, in recent years an alternative interpretation has become increasingly popular. This is the multiverse interpretation of quantum mechanics, initiated by Everett [2]. According to this interpretation, each outcome of a measurement of a variable represents a real development. And the fact that a measurement has produced a certain value just indicates that what the observer thinks of as his universe is nothing more than a branch of the multiverse where the result of the measurement is the true value of the given variable.

Although fundamentally different in the way they perceive reality, it is often supposed that these two views are equivalent for practical reasoning, thus making it more or less a matter of taste which view-point to choose. In my opinion this is a mistake, and it is part of the raison d'être for this paper to argue that the multiverse interpretation can actually be used to explain physical properties that

can not be explained by the Copenhagen interpretation. In particular, this may be the case with the second law of thermodynamics.

In my interpretation, the underlying principle which gives sense to the multiverse interpretation is a property of the dynamical laws which makes equally good sense in a classical setting as in a quantum mechanical one. The essentially classical dynamical law in Section 5 below can be seen as a simple example of this. Thus, different developments which agreed in the past may slide apart just as in the quantum mechanical case.

But there is of course one very important difference: In the classical setting, each possible development will be completely independent of any other, and the lack of determinacy in the dynamical laws will, to an observer living in such a world, appear to be completely at random. In other words, to such an observer, the randomness that he experiences will really look as if God is playing dice. In the quantum mechanical setting however, different developments actually do interact with each other on the micro-level. This is for example the content of the optical experiment where photons pass through two parallel slits, and where the interference that we observe can only be understood if we assume that each photon simultaneously must pass through both slits. Thus an observer can actually perform measurements which confirm the existence of multiple developments.

Einstein's original statement, that "God does not play dice" expressed his deep scepticism towards the probabilistic interpretation of quantum mechanics. In fact, for an almighty God to include randomness in his creation could almost appear as a kind of cheating. However, in the multiverse interpretation of quantum mechanics, nothing happens by chance. Everything which is possible actually does happen, although we can not in general experience it all simultaneously. Thus in a sense it is rather quantum mechanics, not classical mechanics, which tells us that the world is not probabilistic. Or to put it differently: Quantum mechanics may be the best argument we have for saying that God is not playing dice

Before we proceed to investigate the asymmetry of time, there is also another point which has to be clarified. There are in fact different ways of viewing the arrow of time. The concept which is based on entropy, is often called *the thermodynamic arrow of time*. But an equally natural and important concept is what could be called *the historical arrow of time*. According to this view, what is characteristic of the past is that it is unique and that we can actually remember it. We have in general a very good view of the historical development that has lead us to the state that we observe now, but it is much harder to be sure about the consequences that the present state will lead to. Stated somewhat differently, the past is the direction in which the development is uniquely traceable and there seems to be a unique chain of macroscopic states connecting our present reality back to the Big Bang.

All experience that we have seems to support the belief that these different aspects of time's arrow are equivalent. But even if so, the equivalence is by no means a trivial one. And, especially in situations where our usual intuition does not apply, one should be very careful when identifying them. This is in particular the case when we deal with the state of our universe immediately after the Big Bang (or before a possible Big Crunch). Nevertheless, in this paper I will mostly take for granted that these different concepts of time are equivalent, except possibly in extreme states close to the beginning or end of time.

Let us summarise the idea of the historical arrow of time:

Claim 1 (about the unique past)**.** *Every state that we have ever observed seems to have a unique past.*

In fact, no one has ever successfully tried to explain our present state by assuming both that Caesar did cross the Rubicon and that he did not.

Remark 1. Clearly this is in a certain sense no longer true in quantum physics: The microscopic world is full of counterexamples, e.g. the above mentioned experiment with photons passing simultaneously through parallel slits and afterwards interacting with themselves. And the distinction between macroscopic and microscopic is not always easy to make as testified by Schrödinger's famous cat.

One could wonder if it is a problem in the present situation that Claim 1 may fail at the quantum level. I would argue that it is not. In fact, the Claim *is* of course true in an appropriate sense, i.e. even from the quantum mechanical point of view every state that we have ever observed does seem to have a unique macroscopic history. See however the discussion in Section 7.

If we accept the multiverse point of view, we can add to the Claim above a second one:

Claim 2 (about the non-unique future)**.** *Every state that we have ever observed has a non-unique future.*

To put it in another way: There are, in any possible physical situation, phenomena which are genuinely unpredictable like for instance radioactive decay. According to the multiverse point of view, every such event represents a fork in the road towards the future.

These two claims summarise the basic properties of nature that our dynamical laws should reflect. In the next Section I will discuss in more detail what kind of laws this could imply.

4. The dynamical laws

From my point of view a very natural starting point for trying to understand the obvious time-asymmetry in the two Claims 1 and 2 is the "Principle of Least Action", found in every comprehensive textbook on mechanics. This principle has a long history, but its formulation is generally attributed to Maupertius in 1746. It states that a physical system must develop in such a way that the action of the system is (locally) minimised. The principle is very general and in a certain sense the dynamical laws of both general relativity and quantum mechanics can be formulated in such a way. However, from the point of view of working physicists, this principle is usually considered merely as a convenient tool for obtaining the equations of motion. These usually turn out to be completely deterministic: Given positions and momenta for a system of particles (or the corresponding data for the wave-function) on a certain space-like hyper-surface, the equations of motion can be used to compute the state of the system at any time in the future or in the past, at least in theory.

Although the two ways of looking at the problem, as a minimising problem or an initial-value problem, are usually considered to be perfectly equivalent, let me note that from a philosophical point of view they differ fundamentally in a way which is closely related to the Claims 1 and 2 above, and as a consequence also to the difference between the Copenhagen interpretation and the multiverse interpretation

of quantum mechanics. Whereas the initial-value approach uses only data given at a certain time to compute the future development, the minimising approach uses data at two different moments of time to calculate the development in between. To further stress this point, let me formulate a possible interpretation of this in a way which connects the principle of least action with the two Claims above, and which at the same time clarifies the role of the boundary conditions for the universe:

Principle 1 (of Least Action)**.** *Given the states of a system at two different space-like hyper-surfaces, the development of the system in between is (generically) completely unique (and determined by locally minimising the action).*

Principle 2 (of Multiple Futures (and Pasts))**.** *Given only the state of a system at one space-like hyper-surface, the development of the system is in general* not *unique, neither forwards nor backwards.*

Remark 2. Formulated in this way, these principles are clearly symmetric with respect to the direction of time.

One should be a little bit careful when using the word deterministic in this context. In fact, the multiverse interpretation of quantum mechanics is usually thought of as a completely deterministic theory, since time-translation is a unitary operation. This however, does not mean that the world is predictable to an experimenter confined to one universe.

It should also be noted that these two principles may be said to have different status. The first one is a generally excepted fact. As for the second one, it may very well be controversial. Nevertheless, if we insist on both the symmetry of dynamical laws and the multiverse point of view, this is what we have to accept.

Clearly, if we accept Principle 2, then our assumptions about the boundary conditions in the past and in the future will play symmetrical and equally important roles.

5. Symmetry breaking in a simple dynamical principle

So what kind of symmetric law of physics could possibly generate this manifestly asymmetric behaviour in Claims 1 and 2? In this section I will present a very simple example where what from a deterministic point of view appears to be a time-symmetric dynamical law, from the multiverse point of view gives rise to symmetry-breaking.

Consider a single particle in Euclidian 3-space. In the following (and for the rest of this paper), we will only be concerned with discrete time. Thus the movement of the particle from $\mathbf{a}$ to $\mathbf{b}$ may be described as a sequence of points $\{\mathbf{r}(t_k)\}_{k=0}^{m}$ where $\mathbf{r}(t_0) = \mathbf{a}$, $\mathbf{r}(t_m) = \mathbf{b}$, and where $\{t_k\}_{k=0}^{m}$ is the corresponding sequence of (equidistant) moments of time. To keep things as simple as possible, we will assume that the particle is not interacting with anything at all. From a classical point of view such a particle should thus obey Newton's first law

A deterministic time-symmetric microscopic description of Newton's first law could be formulated as follows: If the positions of the particle at times $k-1$ and $k+1$ are $\mathbf{r}(t_{k-1})$ and $\mathbf{r}(t_{k+1})$ respectively, then the position of the particle at time t_k is given by

$$\mathbf{r}(t_k) = \frac{\mathbf{r}(t_{k-1}) + \mathbf{r}(t_{k+1})}{2}. \tag{1}$$

This can be reformulated in the following way: The dynamical law of the particle is given by the condition

$$(2)\quad J_k = 0 \quad \text{for} \quad k = 1, 2, \ldots, m-1 \quad \text{where} \quad J_k = |\mathbf{r}(t_{k-1}) - 2\mathbf{r}(t_k) + \mathbf{r}(t_{k+1})|.$$

A non-deterministic analogue of this would be to attribute the unnormalized probability

$$p_k = \exp\{-\mu J_k^2\} \tag{3}$$

to the position $\mathbf{r}(t_k)$ at t_k, given the corresponding positions at times t_{k-1} and t_{k+1}, where $\mu > 0$ is some number. If we multiply these probabilities together, we get the probability weight

$$p = \prod_{k=1}^{m-1} p_k = \exp\{-\mu \sum_{k=1}^{m-1} J_k^2\} \tag{4}$$

for the path $\{\mathbf{r}(t_k)\}_{k=0}^{m}$, thus obtaining a kind of statistical Ensemble for the set of all possible paths from $\mathbf{a}$ to $\mathbf{b}$.

Remark 3. The choice of the formula for p_k in (3) may look a bit arbitrary but it is not. It can be argued using the central limit theorem that essentially *any* choice of weight at a sufficiently small scale will result in this kind of expression on a larger scale.

Note also that if we can recapture the deterministic behaviour in (1) from the non-deterministic Ensemble in (4) by maximising p. In fact, maximising p is equivalent to minimising the sum

$$\sum_{k=1}^{m-1} J_k^2 \tag{5}$$

which is obviously done by choosing $J_k = 0$ for all $k = 1, 2, \ldots, m-1$.

The interesting point of the model emerges however, when we start to look at the kind of non-deterministic paths which are likely to lead from $\mathbf{a}$ to $\mathbf{b}$. Let us suppose that $\mathbf{a} = \mathbf{b} = \mathbf{0}$. It is then reasonable to ask where the particle is most likely to reach its maximal distance from the origin. To be more precise, we define q_k as the probability (with respect the Ensemble in (4) that the maximal distance occurs at time t_k. Writing for simplicity r_k instead of $r(t_k)$, we have

$$q_k = P(|\mathbf{r}_k| > \max\{|\mathbf{r}_1|, |\mathbf{r}_2|, \ldots |\mathbf{r}_{k-1}|, |\mathbf{r}_{k+1}|, \ldots |\mathbf{r}_{m-1}|\}) \tag{6}$$

$$= \frac{1}{\Sigma} \int_{\Omega_k} \exp\{-\mu \sum_{k=1}^{m-1} J_k^2\} \, d\mathbf{r}_1 d\mathbf{r}_2 \ldots d\mathbf{r}_{m-1}, \tag{7}$$

where $\Omega_k = \{(\mathbf{r}_1, \ldots, \mathbf{r}_{m-1}); |\mathbf{r}_k| > \max\{|\mathbf{r}_1|, |\mathbf{r}_2|, \ldots |\mathbf{r}_{k-1}|, |\mathbf{r}_{k+1}|, \ldots |\mathbf{r}_{m-1}|\}\}$ and Σ is a just the appropriate normalising factor. A natural guess might be to assume the following typical behaviour: After starting out from $\mathbf{0}$, the particle gradually moves away and somewhere in the middle between $t = 0$ and $t = m$ reaches the maximal distance. After this it then gradually returns to $\mathbf{0}$ in a more or less symmetric way. But even if this type of path may generate the largest probability weights in (4), the number of such paths is small compared to the number of paths which so to speak neglect one of the boundary conditions as long as they can. In fact, the computer plot in figure 1 exhibits a behaviour which is quite different

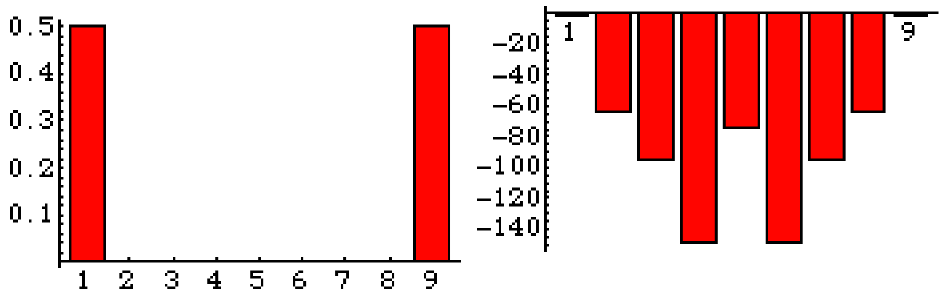

figure 1.

from this (with $\mu = 5$ and $m = 10$). The figure to the left shows the probabilities q_k as a function of time, and the figure to the right shows the same values on a logarithmic scale. Increasing values of μ seem to tend to accentuate the picture. This represents a situation commonly encountered in statistical mechanics where the states with the largest Boltzmann factor $\exp\{-\beta H\}$ are not always the states that occur in reality.

Conclusion 1. *Clearly, we have a break of symmetry: A particle obeying the non-deterministic dynamical principle in (4) will either tend to be close to the deterministic path in the beginning and to deviate in the end or vice versa, but the chance for a symmetric behaviour is very small. This throws new light on the lack of symmetric in Claims 1 and 2. In fact, in view of the very general character of the dynamical principle, one may suspect that this type of behaviour is the rule rather than an exception.*

Remark 4. It should be noted that this phenomenon can also be interpreted as a problem for the use of the calculus of variation on the micro-level: We can not in general in the non-deterministic situation hope that the kind of paths which tend to minimise the expression in (4) (or more generally, the action of a system) will adapt well to the boundary conditions that we may want to impose.

But we can also look upon this from an opposite point of view. It may very well be that a variational principle is the natural way in which nature rules over matter, but that we are actually ourselves interfering with this process by putting up boundary conditions (e.g. by making measurements).

There may be a very interesting link between this and the well-known fact of quantum mechanics that we can not measure the position of a particle without disturbing its dynamical development.

Remark 5. The computer calculations in this paper have all been carried out using Mathematica on a small Macintosh computer. It should be pointed out that the computations are not at all exact, and this is so for very fundamental reasons. When analysing this kind of models, we inevitably come across multiple integrals in a large number of variables and hence encounter what is in the trade sometimes called "The curse of dimensionality problem", i.e. the fact that the time needed to compute multiple integrals grows exponentially with dimension (see [1]). From this point of view, computing integrals with 15 variables is already difficult (except for certain classes of integrands). Using quasi-Monte-Carlo methods (as I do in this paper) one may sometimes be able to obtain some kind of results for, say, 50 or

100 variables, but these should not be confused with exact mathematics. Rather, pictures like the ones in figure 1 should be considered as the results of physical experiments which hopefully, but not with certainty, reflect the true state of affairs.

It is usually not possible to estimate the errors involved, and in some cases one can get results which are obviously not true. Having said this, I want to add that the computations in this paper seem to be relatively stable under changes of the parameters involved which is usually a good indication.

6. The cosmological view of the multiverse

In this section I will introduce a very simplified but full-scale model for the multiverse. Our knowledge of the actual global structure of space-time is limited. We have a reasonably good picture of the "Big Bang" some 14 billion years ago, but we know considerably less about the future. Although many votes in recent years have been in favour of an ever expanding universe instead of a "Big Crunch", we do not really know anything for sure as long as the origin of dark energy, mass and gravitation in general is not well understood.

I have chosen in this paper to work with a bounded model for space-time. The reason is not that this should necessarily be the only or the most likely model. Rather, the motivation is that in this case, the underlying global structure of space-time is symmetric, which makes the problem with the direction of time come out more clearly. Also, the mathematics is simpler in the bounded case, and for the purpose of this paper there is really no reason to choose a more complicated model than necessary.

Remark 6. Before we go further, it may still be worthwhile to comment briefly on the relevance of this discussion for other non-bounded models. If we first consider the case of an infinite ("ergodic") space which is finite in time, then as long as each part of space only interacts with a bounded part of the whole, I believe that basically the same kind of arguments could be applied. Inevitably, the mathematics would have to be more cumbersome in this case. But on the other hand, the same thing could be said about for instance the extension of quantum mechanics, since global wave-functions would no longer form a Hilbert space. Thus the question of whether such an effort is worthwhile must be resolved in view of all other advantages and drawbacks that such a model may have.

As for the case of a space-time which extends infinitely into the future, the question is somewhat different. From a microscopic point of view there is actually no difference at all. The problem is that if we choose to believe in laws of physics which are (essentially) symmetric with respect to time, then in one way or another we will have to make up our minds about the boundary conditions in the future. Even if the same basic mechanism seems to work equally well in the case of an open universe, we simply know less about the constraints that the future may impose on us in this case, in contrast to the closed universe where there is a natural candidate for the boundary behaviour at the end. As a matter of fact, to assume that an open universe will just go on creating its path into the future without any constraints may be just another one of these anthropic traps, based on our deeply rooted view of the future as something which is created out of the present.

From now on the multiverse will be viewed as the collection of all possible developments from the Big Bang to the Big Crunch within the framework of a certain

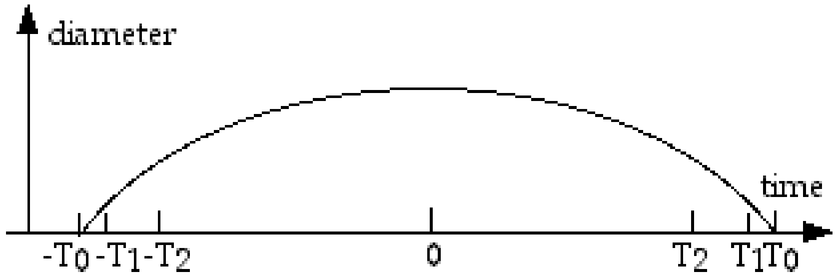

figure 2.

global structure of space-time. This means that we will discard important effects of general relativity and inflation theory, just as we have already decided to discard much of quantum theory. In fact, the kind of multiverse that we will consider will be rather similar to some kind of classical weakly interacting gas. This may be a very bad approximation for most cosmological purposes, but can still be good enough for the purpose of explaining the arrow of time.

Thus consider N particles which we for simplicity consider all to be identical. These particles are supposed to inhabit a global geometric structure extended in time between $-T_0$ and T_0. Furthermore, we assume time to be discrete, and by choosing appropriate units we may assume that it only takes integer values. The volume of the multiverse first expands and the contracts in a way which is essentially symmetric with respect to time. The exact form of the volume as a function of time is not very important in the following, but again for the sake of argument, let us assume that the space-time structure behaves like the curve in figure 2, which reaches its maximal value at the mid-point between $-T_0$ and T_0.

Concerning the end-states themselves, it could perhaps be argued that in a model like this the most natural assumption would be to consider them as not only symmetric states, but as states without any internal structure at all (except possibly for such things as total mass-energy). However, the exact nature of these states is not important in the following, so in view of Ockham's razor one might also argue that we should perhaps not suppose anything at all. What *is* important is that at some point $-T_1$ shortly after the Big Bang, the set $\mathcal{F}(-T_1)$ of possible states of the multiverse that can be obtained from the Big Bang corresponds symmetrically to a similar set $\mathcal{F}(T_1)$ of states at a time T_1 shortly before the Big Crunch. This should not be interpreted as implying that the end-states in a particular universe should be similar in any way. Rather, it is a problem for the model that in a realistic treatment including general relativity, there is no way in which we can simultaneously describe these states within any kind of common geometric frame-work. In Section 10, I will come back to this point.

In the following we will often need to count states. It is therefore important to distinguish two different concepts:

Definition 1. An *elementary state* is a possible state of one particle.

The number $E(t)$ of different possible elementary states at a certain time is assumed to be roughly proportional to the volume $V(t)$ of the geometry at that time.

Definition 2. A *global state* (or just simply a state in the following) is a possible configuration of all the N particles.

Within this framework, an enormous amount of different developments are possible, and to each of them the laws of quantum mechanics will attribute a certain probability. To simplify the model even further we will not consider the precise probability for each development but rather simply classify each step as "possible" or "impossible" thus attributing the (unnormalized) probabilities 1 and 0 to them.

Definition 3. A *universe* is a chain of a states within the given global geometry, one state S_t at time t for each t between $-T_0$ and T_0, with the property that the transition between two adjacent states is always possible according to the dynamical laws.

Definition 4. The *multiverse* is the set of all possible universes in the sense of Definition 3.

The multiverse is thus a kind of huge compact (actually finite) probability space where we have assigned equal probability to all elements. The task is now to try to decide in a heuristic way what types of universes are the most common ones in very much the same spirit as one analyses Ensembles in statistical mechanics.

Definition 5. A certain state S_t at a certain time t is said to be *backwards unique* if there is exactly one possible chain of states connecting it back to a state in $\mathcal{F}(-T_1)$. Similarly, a certain state S_t at a certain time t said to be *forwards unique* if there is exactly one possible chain of states connecting forward to a state in $\mathcal{F}(T_1)$.

Remark 7. To say that a state S_t of a universe is backwards unique is thus the same as saying that it has a unique history. Or as saying that the arrow of time points forwards.

As a matter of definition every state in $\mathcal{F}(-T_1)$ is backwards unique and every state in $\mathcal{F}(T_1)$ is forwards unique. But what happens with the arrow of time in-between? A priori it is not at all obvious that there is any sensible definition of time's arrow in-between. But in the following when discussing possible universes, I will restrict myself to such chains of states where every state on the way from $-T_1$ to T_1 is either backwards or forwards unique. In other words, the only type of universes that will be considered are those where the interval $[-T_1, T_1]$ splits into two parts, $[-T_1, Q]$ and $[Q, T_1]$, where time's arrow points forwards on the first part and backwards on the second part.

Definition 6. The time Q will be called the *peak point* and the time intervals $[-T_1, Q]$ and $[Q, T_1]$ will be called the *backwards* and *forwards unique phases* respectively.

Remark 8. The name peak point of course refers to the fact that from the thermodynamic point of view, this should be the point where the entropy is maximal (see figure 3).

The motivation for assuming that there is a peak point in every universe is not a physical one. Rather, the reason for introducing it is connected to the purpose of this paper. My ambition is to try to argue that the asymmetry of time in our universe is very well compatible with a multiverse governed by completely symmetric principles, so I will be content with comparing developments which are sufficiently general to illustrate this point. Thus the question why we live in a world where it makes sense to speak of history at all, and likewise the question how the switching from leftward to rightward uniqueness actually takes place, are somewhat outside the scope of this investigation.

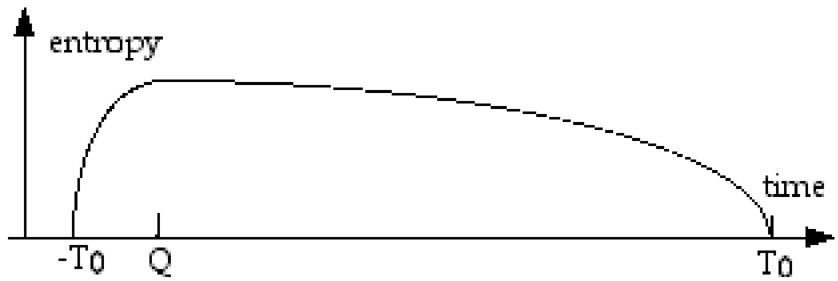

figure 3.

Now consider a certain state S_t at a time $t \in [-T_1, Q[$. Thus S_t is backwards unique but according to the discussion in Section 3 we do not in general expect it to be forwards unique, So how many different states are possible to obtain at time $t+1$ from S_t? Since an exact computation is clearly impossible to make, the answer will have to be given by a crude approximation:

Heuristic Assumption 1. *For any $t \in]-T_1.T_1[$, the number of states at time $t+1$ obtainable from the given state S_t is given by a constant K which is roughly independent of S_t. Similarly, the number of states at time $t-1$ obtainable from the given state S_t at time t is given by K.*

Remark 9. The underlying reason for this assumption is that as long as the number $E(t)$ of possible states at time t accessible to the particles is much larger than the number N of particles, the number K should depend on N, rather than on $E(t)$. Whether the universe is expanding or contracting, or whether the density is high or low within reasonable limits, we still assume that K essentially depends on the number of choices (or "forks in the road") open to each individual particle. Needless to say, this approximation must be a very bad one in a realistic multiverse. But it could still be reasonable in the kind of situation we are dealing with.

7. Time's arrow, a heuristic approach

Let us now suppose that we start with anyone of the K_1 states in $\mathcal{F}(-T_1)$. According to the heuristic assumption 1, this state will after one unit of time give rise to K possible states, after two units of time we have K^2 states and so on. T units of time after $-T_1$ we will have a total of

$$K_1 K^T \tag{8}$$

possible states. But are they all different? If they are then all the developments leading to them are all backwards unique. On the other hand, clearly when the global geometry of the multiverse starts to shrink and thus the number of states starts to decrease, we must sooner or later reach a point where the actual number of states will be less than the number of states predicted by formula (8). But then of course, not all of them can be backwards unique, since some of them will have to end up at the same state.

Whether this will happen or not before time T_1 clearly depends on the size of the number K. And to estimate this accurately seems difficult. On the other hand, as stated in Claim 1, all experience that we have from our own universe definitely tells us that history tends to be unique. We will return to this question in Remark 12 below.

Heuristic Assumption 2. *When computing the chance that different developments actually will meet at a certain time t, we will in the following assume that they behave as statistically independent variables.*

Remark 10. This assumption must be interpreted with some care. In fact, quantum mechanics tells us that developments which have very recently separated have a good chance of interfering with each other shortly afterwards. In the following I will assume that the constant K measures the "effective" rate with which the number of the number of possible states grows.

Of course, the above heuristic assumption could not possibly be true in any strict sense in a multiverse inhabited by the kind of particles we are used to. In fact, even in universes with very different developments, the particles would probably tend to organise themselves in similar structures, like atoms and molecules.

Suppose now that we have numbers $T_0 > T_1 > T_2$ such that

$$T_0 - T_1 \ll T_0 - T_2 \ll T_0. \tag{9}$$

As has already been said, the exact form of the dependence of the diameter on t is not at all important, but for the sake of definiteness we may for instance assume that, say, $D(t) = D_0 \cos(t\pi/2T_0)$, which gives a similar formula for the volume $V(t)$ as a function of t: $V(t) = V_0 \cos^3(t\pi/2T_0)$.

Claim 3. *Suppose that K is such that the chance for two different developments starting from some states at time $-T_1$ to meet before time T_1 is neglectably small. Then the number $\mathcal{N}(t)$ of universes with peak-point at $t \in]-T_1, T_1[$ is proportional to .*

$$\frac{1}{G(t)}, \tag{10}$$

where $G(t)$ is the number of (global) states of the global geometry at time t.

The heuristic idea here is very simple: We can compute the number of backwards unique developments up to time t and we also can compute the number of forwards unique developments from t and onwards. To get the total number of developments with peak point at t we need only estimate how many such backwards unique developments will fit onto a forwards unique continuation. This is easy if we make use of the heuristic assumption 2.

In fact, the number of such universes is given by the total number of backwards unique developments up to time t multiplied by the probability that such a development will fit onto a forwards unique development from the future. From the heuristic assumption 2 we see that the latter is given by the number of such possible developments divided by $G(t)$. Thus according to the heuristic assumption 1:

$$\mathcal{N}(t) \propto K_1 K^t \times \frac{K_1 K^{(2T_1 - t)}}{G(t)} = \frac{K_1^2 K^{2T_1}}{G(t)} \propto \frac{1}{G(t)}. \tag{11}$$

Remark 11. Clearly, this way of calculating $\mathcal{N}(t)$ is symmetric with respect to the direction of time.

We have now arrived at the following

Claim 4. *The overwhelming majority of all possible universes have their peak-points in $[-T_0, -T_2]$ or $[T_2, T_0]$, not in $[-T_2, T_2]$.*

In fact, in view of the previous Claim 3, the number $\mathcal{N}_2$ of universes which peak in $[-T_2, T_2]$ is given by

$$\mathcal{N}_2 \propto \sum_{t=-T_2}^{T_2} \frac{1}{G(t)}. \tag{12}$$

The particles are supposed to be identical, so the number $G(t)$ is given by

$$G(t) = \binom{E(t)}{N} \approx \left(\frac{eE(t)}{N}\right)^N. \tag{13}$$

Since we have assumed that $E(t) \propto V(t)$ and $V(t) = V_0 \cos^3(t\pi/2T_0)$, we readily obtain

$$\mathcal{N}_2 \propto \sum_{t=-T_2}^{T_2} \frac{1}{\cos^{3N}(t\pi/2T_0)} \propto (T_0 - T_2)^{-3N+1}. \tag{14}$$

On the other hand, the number of universes which peak in $[-T_1, T_1]$ is by a similar calculation:

$$\mathcal{N}_1 \propto \sum_{t=-T_1}^{T_1} \frac{1}{\cos^{3N}(t\pi/2T_0)} \propto (T_0 - T_1)^{-3N+1}. \tag{15}$$

Thus, since the ratio between the number of developments with peak point in $[-T_2, T_2]$ and the total number of developments is obviously less then the corresponding ratio between the number of developments with peak point in $[-T_2, T_2]$ and those with peak point in $[-T_1, T_1]$, we get the following estimate for the percentage of developments with peak point in $[-T_2, T_2]$,

$$< \frac{(T_0 - T_2)^{-3N+1}}{(T_0 - T_1)^{-3N+1}} = \left(\frac{T_0 - T_1}{T_0 - T_2}\right)^{3N-1}, \tag{16}$$

which will be extremely small in view of (9). This proves the claim.

8. A numerical comparison

The above construction could at best be said to model a kind of idealised semi-classical multiverse inhabited by identical particles which behave more or less like a weakly interacting gas. Although I make no claims what so ever for this model to reflect the details of our actual multiverse, it can still be interesting to insert some numerical data from cosmology to show what kind of magnitudes are involved if we consider a structure of the size of our observable universe.

Thus let the number of particles N be of the order of magnitude (corresponding to the number of particles in our observable universe):

$$N \sim 10^{80}. \tag{17}$$

We take the basic unit of time to be the Planck time $= 10^{-45}$ seconds. Suppose also (quite arbitrarily but for the sake of argument) that the life-span $M = 2T_0$ of this multiverse is of the order of magnitude

$$M \sim 10^{64} \quad \text{in Planck time units} \tag{18}$$

which approximately corresponds to 300 billion years. The number of possible different elementary states within a volume V that a particle can occupy is (again quite arbitrarily) taken to be approximately equal to the number of cells with

diameter equal to the Planck length ($\sim 10^{-33}$ cm) that it contains, which gives approximately 10^{99} states/cm^3.

The difference $T_0 - T_1$ must be chosen so large that the number of states at time T_1 is much larger than the number of developments leading there. If we for instance take $T_0 - T_1 = 10^{35}$ ($= 10^{-10}$ seconds), then according to standard inflation theory the diameter of the global geometry may be of the order of magnitude $\sim 10^{12}$ cm, which gives a total of 10^{135} elementary states. Thus the number of possible states for the 10^{80} identical particles is

$$\binom{10^{135}}{10^{80}} \sim 10^{\left(10^{82}\right)} \tag{19}$$

If we for example let K be of the order of magnitude

$$K \sim 10^{\left(10^{10}\right)}, \tag{20}$$

then we get for the number of universes (neglecting K_1)

$$K_1 K^M \sim \left(10^{\left(10^{10}\right)}\right)^{10^{64}} = \left(10^{\left(10^{74}\right)}\right) \ll 10^{\left(10^{82}\right)}. \tag{21}$$

If we now set $T_0 - T_2 = 10^{45}$ (one second), then we see from (16) that the fraction of universes with peak points further than one second away from the Big Bang or the Big Crunch is at most of the order of magnitude

$$\left(\frac{T_0 - T_1}{T_0 - T_2}\right)^{3N-1} \sim 10^{-10^{81}}. \tag{22}$$

Remark 12. The reader has probably noticed the arbitrariness in the choice of K. I do not know what a reasonable value would be or how to compute it. However, it should be noted that if we accept the multiverse point of view, then one can argue that it is an experimental fact that K can not be too large. In fact, we have already experienced our universe for something like 10^{63} Planck units of time, without the slightest indication that our present state could be explained in terms of macroscopically different histories.

Thus, if we estimate the diameter of our (observable) universe to something like 10^{32} cm, then computing the number of possible states at present in the same way as before, we see that the fact that our history seems to be unique indicates that

$$K^{\left(10^{63}\right)} \ll \binom{10^{195}}{10^{80}} \sim 10^{\left(10^{82}\right)}. \tag{23}$$

This is in fact not so far from what is needed for the discussion above (compare (21)).

9. The polygonal multiverse

In the previous section I tried to argue on heuristic grounds that the second law of thermodynamics and the arrow of time could be viewed as the result of a break of symmetry in the multiverse. Although the numerical bias in favour of universes with directed time is enormously large, the assumptions leading to this conclusion are so heuristic that it is certainly appropriate to question there value. Therefor, in this section I propose an attack on the problem from the opposite side: Instead of considering the whole multiverse even in the simplified form of the previous section, let us consider very small and simple models for it, and in these models try to prove,

by mathematics or by computer computations, that the kind of symmetry breaking encountered in section 8 actually occurs.

In this paper, I will content myself with discussing one model, where numerical computations are possible and symmetry breaking occurs, but where the number of particles is much to small to speak of entropy in an interesting way and hence where the concept of order/disorder has to be defined in a much more primitive manner.

Thus, let the global structure be given by the set

$$M = S_1 \times \mathbb{N}_M. \tag{24}$$

Here the unit circle $S_1 = \mathbb{R}/\mathbb{Z}$ furnishes the spacial part of the global structure, and $\mathbb{N}_M = \{-m, -m+1, \ldots, -1, 0, 1, \ldots, m-1, m\}$ (where $M = 2m$) is the discrete time ranging from $-T_0 = -m$ to $T_0 = m$ in the terminology of Section 8. We suppose that this global geometry is inhabited by n particles x_j, $j = 0, 1, \ldots, n-1$. A universe in this context is simply a collection of n sequences $\{x_{j.k}\}_{k=-m}^{m}$, where $x_{j.k}$ is the position of the j:th particle at time k. Thus, for each moment of time k we can think of the state of the particles as a n-polygon with corners at the points $x_{0,k}, x_{1,k}, \ldots, x_{n-1,k}$ on S_1.

The radius (and perimeter) of the spacial part varies with the time k as $r_k = \cos(\pi k/m)$. However, it is more convenient to use the same coordinate $x \in [0, 1[$ on $S_1 \times \{k\}$ for all k and instead include the expansion and contraction of space in the dynamics (see (27) below).

At $k = -m$ and $k = m$ ($-T_0$ and T_0), we suppose that these n-polygons are regular with corners at the points $0, 1/n, 2/n, \ldots (n-1)/n$. More precisely, we assume that the j:th particle x_j has coordinate j/n. In a deterministic model, the dynamics of the model would consist in an expansion and a contraction as above with each particle having the same coordinate all the way from $k = -m$ to $k = m$. Thus at each moment of time the corresponding polygons would still be completely regular. This can be thought of as a universe in which the only dynamical principle which comes into play is Newton's first law. It is created perfectly uniform and preserves this property throughout its existence.

But we can now, in analogy with what was done in section 5 also introduce a non-deterministic dynamical principle. In this case we simply assume that each particle x_j interacts with its neighbour x_{j+1} in the following way (we identify x_n with x_0): The most probable distance from x_j to x_{j+1} at time k is the average of the corresponding distances at time $k-1$ and $k+1$, or in other words,

$$x_{j+1,k} - x_{j,k} = \frac{1}{2}((x_{j+1,k-1} - x_{j,k-1}) + (x_{j+1,k+1} - x_{j,k+1})) \tag{25}$$

Hence, if we define

$$J_{j,k} = (x_{j+1,k-1} - x_{j,k-1}) - 2(x_{j+1,k} - x_{j,k}) + (x_{j+1,k+1} - x_{j,k+1}) \tag{26}$$

then the deterministic principle is equivalent to the requirement that $J_{j,k} = 0$ for all $j = 0, 1, \ldots, n-1$ and $k = 1, 2, \ldots m-1$. The reader may find it amusing to think of this law as a kind of Mach's principle in this miniature world. The non-deterministic principle is now obtained by replacing this condition by a probability weight

$$p_{j,k} = \exp\{-c_k J_{j,k}^2\}, \tag{27}$$

where the c_k:s are constants. Of course, we want to assume the dynamical law to be the same all over, but here we must also remember that the length-scale on the spacial part S_1 varies with time. Hence, the coordinates $x_{j,k}$ should actually be multiplied with r_k to give the correct length. For this reason we will put $c_k = c(r_k)^2$. This gives a slight bias in weighting between the influence from $k-1$ and $k+1$ respectively when the radius of the spatial part expands or contracts, which I however will not consider important.

Thus, the unnormalized probability for a given universe is

$$p = \prod_{j,k} p_{j,k} = \exp\{-c_k \sum_{j,k} J_{j,k}^2\}, \tag{28}$$

which gives us the Ensemble.

Remark 13. Clearly both the Ensemble and the deterministic law (25) are time-symmetric, i.e. remain unchanged under the transformation $k \mapsto -k$.

From this Ensemble we can clearly recapture the deterministic dynamical law in (25) above by maximising p, which is equivalent to minimising

$$\sum_{j,k} J_{j,k}^2. \tag{29}$$

This is clearly done (as in Section 5) by choosing all $J_{j,k} = 0$.

We will only be interested in values of c which are comparatively large, in which case the developments which have a reasonable chance of occurring can be considered to be small perturbations of the deterministic one. In this case, the chance that two particles will ever cross each others paths is very small, and it is obviously meaningless to talk about entropy of the polygon in any usual sense. But we can instead measure how irregular it is at time k by measuring how much its centre angles differ from the regular value $2\pi/n$. Since the centre angles are given by $2\pi(x^{j+1} - x^j)$, we can thus take

$$S_k = 2\pi \sum_{j=1}^{n} |x_k^{j+1} - x_k^j - 1/n| \tag{30}$$

as a measure of irregularity or "disorder" in the polygonal universe at time k. Note that this measure only measures the form of the polygon and is independent of the actual size of the global geometry.

Figure 4 shows a computer plot (using Mathematica) of the probability q_k that S_k at time k is larger than the deviation at every other time, and in figure 5 we have the corresponding logarithmic plot. In this picture, the parameters have been chosen to be $M = 2m = 30, n = 3$ and $c = 25$. Clearly, in this case we see that with very high probability, maximal disorder occurs close to one of the end-points $-T_0$ and T_0.

10. Time's arrow, conclusions

I this paper, I have discussed the arrow of time from different perspectives. In a way, the results that come out all point in the same direction: Whenever deterministic dynamics is replaced by the multiverse point of view, the symmetry of time tends to break in favour of a directed time. In the global setting, the consequence

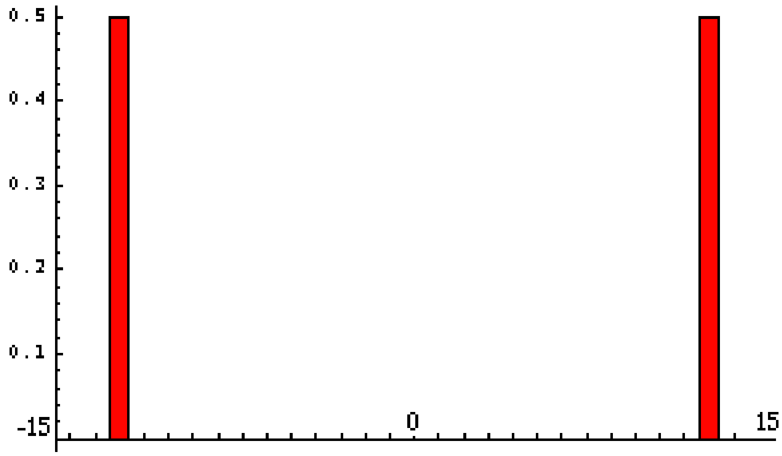

figure 4.

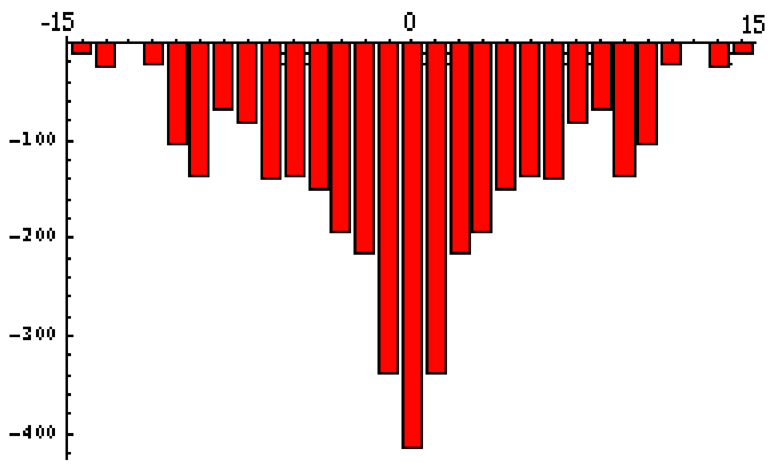

figure 5.

is that the arrow of time may point in different directions in different parallel universes: What is the future to us may very well be history to the inhabitants of another parallel universe.

However, the gap between the cosmological model in Sections 6-8 and the micro-multiverse in Section 9 is too large to be satisfactory. What I would really want to do to bridge this gap is to apply the computational approach of Section 9 to a model which is large enough to make real computations with entropy meaningful. A nice example would perhaps be furnished by a classical gas of, say, 100 particles in three-space, weakly interacting in agreement with some dynamical principle, on a time-axis from $-T_0$ to T_0, containing perhaps 100 units of time. This would probably be enough in order to study the entropy and to look for conditions on the dynamics which would assure that the entropy with overwhelming probability assumes its maximum close to one of the end-points. However, this problem may be very difficult in view of the computational difficulties involved. In fact, numerical computations in such a model would involve integrating 30000 variables. In view of

"the curse of dimensionality" discussed in Remark 5, this is clearly far out of reach for the somewhat primitive computational methods used in this paper. In other words, to master such a model seems to be an interesting and highly non-trivial problem for research, the solution of which will probably have to make use of both computational power and clever mathematics.

But even if this could be managed, how relevant is a model which behaves essentially like some kind of weakly interacting classical gas in an essentially Euclidian global geometry with finite volume, no curvature and no inflation? To accept Euclidian geometry and a classical approximation to quantum mechanics for the purpose of studying the second law of thermodynamics can in my opinion be rather reasonable, the weakest part of it being the schematic way of classifying developments simply as possible and impossible as was done in Section 6 instead of trying to compute the actual probabilities. As a matter of fact, at least from my point of view, entropy can be thought of as an essentially classical macroscopic concept. More problematic points are:

(1) The idea of the particles as weakly interacting.
(2) The behaviour near the end-points.
(3) The behaviour in the future if the global geometry will expand for ever.

The particles in our universe are definitely not weakly interacting but rather tend to group themselves together into things like atoms, human beings, planets and stars. This will of course very much influence the behaviour of the entropy. But in this sense, probably every model that we will ever construct will have to be a coarse approximation, even if one can hope to do much better than in this paper.

As for the behaviour near the end-points, something better should definitely be done. In Section 6, I have eliminated this behaviour simply by starting the analysis at times $-T_1$ and T_1 instead of $-T_0$ and T_0, and by requiring there to be a symmetry between the sets of possible states at these times. It is not necessarily a problem in itself that these sets may contain states of very different character. But it is definitely unrealistic to think that these states could be described by the same global geometry: A Big Crunch with enormous black holes somehow being squeezed together into zero volume must surely have a geometry very different from the smooth Big Bang. At this point we must leave the classical approximation and involve both general relativity and quantum theory. And it may be that something similar to Hawking's analysis mentioned in Section 2 could do the job.

The case of an ever expanding global geometry may be seen both as a problem and as a possibility. It can be seen by computer experiments that the kind of behaviour that we have encountered earlier can occur also without assuming a Big Bang and a Big Crunch. I have not included any extra figures to illustrate this point, but figure 1 in section 5 can actually be interpreted as representing a broken symmetry in a multiverse with constant volume and only one particle.

The real problem here is that we do have to make up our minds about the "boundary condition" or more generally about the behaviour in the future. As an example, I can mention that if we consider an expanding model with a dynamical principle of the type we have met in Sections 5 and 9, but put a perfectly ordered boundary condition at the expanding end (for practical reasons after a finite time), then we get a multiverse where the peak point is extremely likely to lie close to the starting point: In other words, from the point of view of a possible inhabitant of

such a universe, history would start from an ordered state of infinite volume and then steadily contract into a Big Crunch.

One way out is of course to adopt the view that the explanation of the asymmetry of time is to be found in a corresponding asymmetry in the underlying laws of nature. In this case it may be that the relevant microscopic dynamical principle will only depend on information from one direction. If the final cosmological conclusion will be that our universe is expanding for ever, then this may in fact in a way be the most attractive solution.

But if we insist that the underlying dynamical laws should be (essentially) time-symmetric, then we must inevitably say something about the future since in this case it would be no more logical to say that the future does not influence us than to say that the past does not.

As I see it, there are in this case two ways out: Either we try to find a boundary condition at infinity which in some sense we can argue is natural. Or we can argue that most of the possible boundary conditions should be very disordered for statistical reasons. The sad part is that it is so much more difficult to make observations of the future: It could be in this case that we would simply have to live with our ignorance.

Another point that I would like to comment is the question how natural the dynamical principles of Sections 5 and 9 are. One could of course say that physicists have a long-time habit of constructing variational principles to fit with practical needs. But if one really believes that Maupertius idea contains some fundamental truth of nature, this argument is clearly not good enough.

The argument that I have already mentioned in Section 6 is better: Even if the construction process is not very natural, the central limit theorem shows that the final result is more or less independent of what kind of stochastic process is involved, as long as all this takes place on a much smaller scale.

However, from my own personal point of view neither of these motives were decisive. As a matter of fact, my interest in this field started within general relativity, where I from a strictly geometric point of view have used the idea of multiple histories to derive a kind of geometric principle of least action (see [6]). The dynamical principles in this paper can in a certain sense be viewed as simple analogues of that. This geometric principle of least action is defined geometrically in terms of the curvature tensor and does not depend on time in any other way than through the metric.

References

[1] Carvajal, O. A., Chapman, F. W. and Geddes, K. O., Hybrid Symbolic-Numeric Integration in Multiple Dimensions via Tensor-Product Series. Preprint. Revolution. London: Viking.
[2] Everett, H., 'Relative State* Formulation of Quantum Mechanics. Reviews of Modern Physics, 29, 1957.
[3] Gold, T., The Arrow of Time. American Journal of Physics, 30, 1962.
[4] Hawking, S. W., Arrow of Time in Cosmology. Physical Review D, 33, 1985.
[5] Penrose, R., Singularities and Time-Asymmetry. General Relativity: An Einstein Centenary, Cambridge University Press, 1979. Oxford, University Press, 1996.
[6] Tamm, M., A Statistical Approach to the Concept of Mass. Physics Essays Vol 18.2, 2005. Ultimate Reality. Cambridge University Press, 2003.